DOCUMENT RESUME

ED 235 959 RC 014 398

AUTHOR McConnell, Beverly B.
TITLE Evaluation of Computer Assisted Instruction in Math, Pasco School District. Final Report.
INSTITUTION Pasco School District 1, Wash.
PUB DATE Sep 83
NOTE 29p.; Study designed by John Mattson.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Computation; *Computer Assisted Instruction; Elementary Education; *Elementary School Mathematics; Grade 3; Grade 6; Limited English Speaking; Mathematical Concepts; *Mathematics Achievement; Mathematics Curriculum; *Mathematics Instruction; *Mathematics Skills; Migrant Children; Remedial Instruction; Special Education
IDENTIFIERS *Pasco School District WA

ABSTRACT

Computer assisted instruction (CAI) significantly improved the math skills of elementary school children, according to the results of an experiment conducted by the Pasco (Washington) School District. The experiment was designed to determine whether CAI would improve student math performance and whether CAI would improve math performance more than a less expensive alternative program. During two semesters, approximately 500 students in grades 3-6 received one of three math treatments: CAI, the alternative paper and pencil drill and practice program, or the regular district math curriculum. While all three treatments brought children from below to above age grade norms in total math, CAI improved the learning of total math and computational skills significantly more than the other two treatments. The regular curriculum was more successful in improving concept application skills. CAI was most effective at the third grade level and least effective at the sixth. CAI was also very effective for children enrolled in special education and remedial programs, for migrant children, and for children with limited English proficiency. A positive correlation existed between length of time spent on the computer and gain in math performance. A high and consistent correlation also existed between grade equivalent ratings on standardized tests and those ratings derived from the computer. (SB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED235959

EVALUATION OF COMPUTER ASSISTED INSTRUCTION IN MATH

PASCO SCHOOL DISTRICT

STUDY DESIGNED BY JOHN MATTSON

DATA ANALYSIS AND REPORTING

BY

BEVERLY B. MCCONNELL, Ph.D

FINAL REPORT

September, 1983

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

RC014398

PASCO SCHOOL DISTRICT NO. 1

1004 NORTH SIXTEENTH PASCO, WASHINGTON 99301

PHONE (509) 547-9531

EXECUTIVE SUMMARY

This report is an analysis of math scores by children in the third through sixth grades who were receiving one of three math treatments; computer assisted instruction (CAI), an alternative paper and pencil drill and practice program, or just the regular district math curriculum. The experiment took place at three schools within the Pasco School District and was designed to answer the question: Will computer assisted instruction significantly improve children's math performance?

The primary findings are as follows:

- The gains made by children in CAI were consistently higher in computation and total math than gains made by children in either of the other math treatments. This superiority was large enough to be statistically significant.
- Gains by children in CAI were much greater than "expected" gains based on national norms. The rate of acceleration of learning was large enough to exceed the standard set by the U.S. Dept. of Education as "educationally significant."
- The CAI program produced the greatest gains in computation skills and total math. It was less effective than the district math program in producing gains in concepts application.
- All special groups of children for whom separate analysis was done made more gains under CAI than under the other two treatments. This included children in special education, children eligible for remedial assistance in Chapter 1 or RAP, migrant children and children of limited English proficiency.
- By the standardized tests, children in third grade benefitted by the CAI program the most, sixth grade children the least. However, by the grade equivalent gains reported by the computer, geared to the exact skills taught in the computer program, sixth grade children made some of the largest gains. This finding is therefore probably an indication of a poor match between the curriculum and the content of the standardized test.

DESCRIPTION OF THE PROGRAM

In the Spring of 1982 Pasco School District received funding for five computer terminals. These were installed in a classroom in Longfellow Elementary School. A software program for elementary math skills was contracted for from a terminal in Portland, Oregon. A brief trial run use of the terminals was carried out in the Spring of 1982 to work out the logistics of how children could be scheduled in, five at a time, for ten minute sessions on the terminal, released from regular classes.

A paraprofessional teacher was hired and given one day training on how to manage the program. Children from the third, fourth, fifth and sixth grades at Longfellow schools were randomly assigned to two groups. One group received computer assisted instruction during the fall semester in 1982 while the other group were given a paper and pencil drill and practice program in their classrooms each day for an equivalent period of time. These practice sheets were purchased from the same publisher from which the district math workbooks were obtained and were meant to reinforce math skills taught in the district math program. The second semester in the Spring of 1983 the groups were reversed. Children who had been receiving CAI (computer assisted instruction) did the drill and practice worksheets and the second group was assigned to work on the computer terminals.

The manager for the CAI program reported that children were able to learn to operate the terminals with only a few minutes coaching. At the beginning of the semester ten days of placement programming is carried out in which the children are given different

types of math problems at different levels, and their individual strengths and weaknesses are assessed. The basic concept of the math skills had presumably been learned by the children in regular classes, and the computer work was intended to reinforce skills already learned rather than to provide initial instruction. Following the placement period, an individual program was established through the computer for each child emphasizing areas of weakness in math skills as shown by the assessment of individual performance. There were fourteen tracks of math skills taught. The computer did not assign problems in tracks which the child had already mastered. Also when a child completed a particular math track, the record would show that this skill area had "topped out" and the computer would assign work in other tracks.

When a problem appeared on the terminal, the child had 30 seconds to type in the correct answer. If a wrong answer was typed in,the computer signified "wrong" and presented the problem again. At the end of 60 seconds after presenting a problem, if there was no answer by the child, the computer would tell the child what answer to type in and, after providing the correct answer, the child would go on to the next problem. The computer would record each child's performance and would score a problem as "wrong" if the child had given the wrong answer or if it took more than 30 seconds to provide the right answer.

At the end of an individual session the computer would report the number of problems attempted, the number and percentage correct. The lab manager would record this information and use it to

reinforce children for their efforts by means of a "star chart" in which a child received stars of different colors for correct performance of over 70%, over 80%, over 90% or 100%. The lab manager awarded special recognition for performance of 100% correct, as well as for consistent performance of 90% or better. She reported that the children took great interest in the star charts and appeared to be motivated by them.

At the end of each week, the school received from the parent terminal a report in which a child was given a grade equivalent score in whichever math track the child had worked on that week. The report also provided cumulative grade equivalent score (across all tracks) which provided a rough measure of a child's progress. Copies of these reports were given to the classroom teachers. No internal use was made of the reports for planning instruction because the computer did this as part of the basic program sequence.

Children were assigned to the computer room approximately ten minutes a day. If a child was absent, an effort was made to make up the missed sessions. Overall the range of time children had on the terminal ranged from six to thirteen hours, with the majority of children receiving from ten to eleven hours, or 55 to 65 ten minute sessions over the course of a semester.

A control group of children who did not receive extra reinforcement of math skills by either the drill and practice worksheets, or the computer assisted instruction, was tested in two neighboring schools. These were children in the regular district math series. The schools selected were those known by the district

evaluation staff to have children comparable in test performance on other district tests to those enrolled in Longfellow School.

THE EVALUATION DESIGN

The experiment was designed to answer the question: Will computer assisted instruction significantly improve children's math performance? Will it improve math performance more than a less expensive alternative "drill and practice" program?

There were many side questions to these basic questions: Will CAI have different effects in reference to different types of math skills? Will CAI have different effects for different student populations--particularly those eligible for various types of special services such as migrant, remedial students eligible for Chapter One or district remedial funds, limited English speaking children, and children in special education.

Children were tested in the fall, mid-year, and spring. Third, fifth and sixth grade students were given the CTBS (Comprehensive Test of Basic Skills, Form U published by CTB/McGraw Hill in 1981). Fourth grade students were tested with the CAT (California Achievement Test, Form C, published by CTB/McGraw Hill in 1977). It is not ideal to have two different tests used, especially when test results are to be combined to assess a program wide effect. However the use of two tests was necessary because the CAT is mandated by the state for fourth grade testing state-wide, and it would have been an added expense for the district to use another test, and would have taken up student and teacher time to administer

another test. To overcome problems of combining scores from two tests, the analysis used a metric common to both tests, the normal curve equivalent, or NCE. And the content of both tests was parallel in the subject areas of the subtests in math.

Tests were administered by regular classroom teachers. Very little training was given in testing procedures; however most teachers had some previous testing experience. Test score conversions were made by district staff in each school and provided to the evaluator. Before the final analysis an accuracy recheck system was used by the evaluator to validate test score conversions, and corrections were made before putting the completed test data on computer for analysis.

Each of the tests used had norms for fall, mid-year and spring testing. These were used for score conversions appropriate to the test data.

The metric used for the evaluation is NCE scores (normal curve equivalent). These are scores which can be used for statistical analysis because they represent an equal interval scale--i.e. a change of 2 points for a child at one end of the spectrum of scores is supposedly the equivalent of a 2 point change for another child at a lower point in the spectrum. Because of this characteristic, it is possible to combine the scores of children at different grade levels, and using different tests, to obtain a measure of comparative change for the program as a whole.

In an evaluation using NCE's, an estimate of "expected performance" would be that the child would have the same score in mid-year or in the spring as in the fall since the score measures

the performance of the child in reference to other children in the norm group. To obtain the "same" NCE at subsequent testings, a child must make an improvement in raw score since all children in the norm group will also be making some improvement based on their schooling and added maturity. If the child has an NCE of 40 in the fall and again at mid-year, this means that the child is 10 points below the norm (which in NCE's is 50 for each grade level) but is improving at the same rate as other children, with the result that the relative performance at mid-year is the same. If the child shows a gain in NCE score, it means that relative to the norm group the child is gaining--making a faster than average improvement in skills.

The National Department of Education considers a gain of 3 NCE's an indication of a "successful" remedial effort. A gain of 7 NCE's (in a school year) is considered an "educationally significant gain which is the standard used in judging "exemplary" programs. This provides a useful "rule of thumb" as to whether a particular program is as successful or more successful an effort to improve academic performance than other experimental programs nationwide.

Between the experimental and control group schools, there were 557 children tested in the fall of 1982, at the beginning of the experimental period. Because of children moving away or absences on testing dates, there were 502 of these children who completed tests at mid-year (90%), and 429 tested again in May, 1983 (77%) at the end of the school year. This represents a relatively high retention rate for an evaluation of an educational

program, and a large enough group at each grade level to provide stability of statistical analysis.

A separate analysis was done of children classified as special education students, and the scores of these children were not included in the overall program analysis. Separate analysis was also done for children eligible for various special funding categories--remedial, migrant, or limited English proficiency. All of these groups, however, were also included in the overall analysis of program results. The numbers of children in some of these special categories was quite small so that statistical analysis may be subject to more variability because unusual performance by one or two students can effect group averages more when numbers are small. Findings, therefore, must be reviewed with caution for the effects of the program in these subgroup analyses.

All data was coded for computer analysis, using the SPSS system available through the computer center at Washington State University.

FINDINGS

Were the treatment and control groups comparable before the experiment?

Yes. Statistical analysis of the fall "pre-tests" for all three treatment groups (CAI, Drill and Practice, and Control) showed the differences between the groups were not statistically significant either for the total math score or for the subtests of computation and concept application. This was true in the analysis of the total program, and in the separate analysis by treatment groups at each grade level. Overall all three treatment

groups started the program year with average scores which were below national norms for children of comparable grade levels, and these scores were not significantly different from one another in the three groups. Details of this statistical analysis are given in Table 1 later in this report.

Did computer assisted instruction, CAI, improve the learning of math skills more than other treatments?

The answer to this question appears to be "yes." The gain scores for children in the fall and the spring semesters is presented in Figure 1 which shows gains separately for the three different types of math program treatment.

As can be seen in the fall semester the children receiving computer assisted instruction, CAI, made larger gains than the control group at every grade level, in the total math score and in the subtests of computation and concepts application. The CAI treatment appears to have been most helpful to improving skills in computation skills where the gains are much larger than in the area of concepts application. It also seems to have been more effective with the younger children; the largest gains are at third grade. The smallest gains are at the sixth grade level where the scores were only marginally better than the control group, and somewhat less than the children in the alternative drill and practice treatment.

The statistical analysis shown at the bottom of each column of gain scores shows that the differences between groups were statistically significant in computation, concepts application, and in total math for the fall semester. "Significant at the .001 level" as shown on the figure translates to mean that the

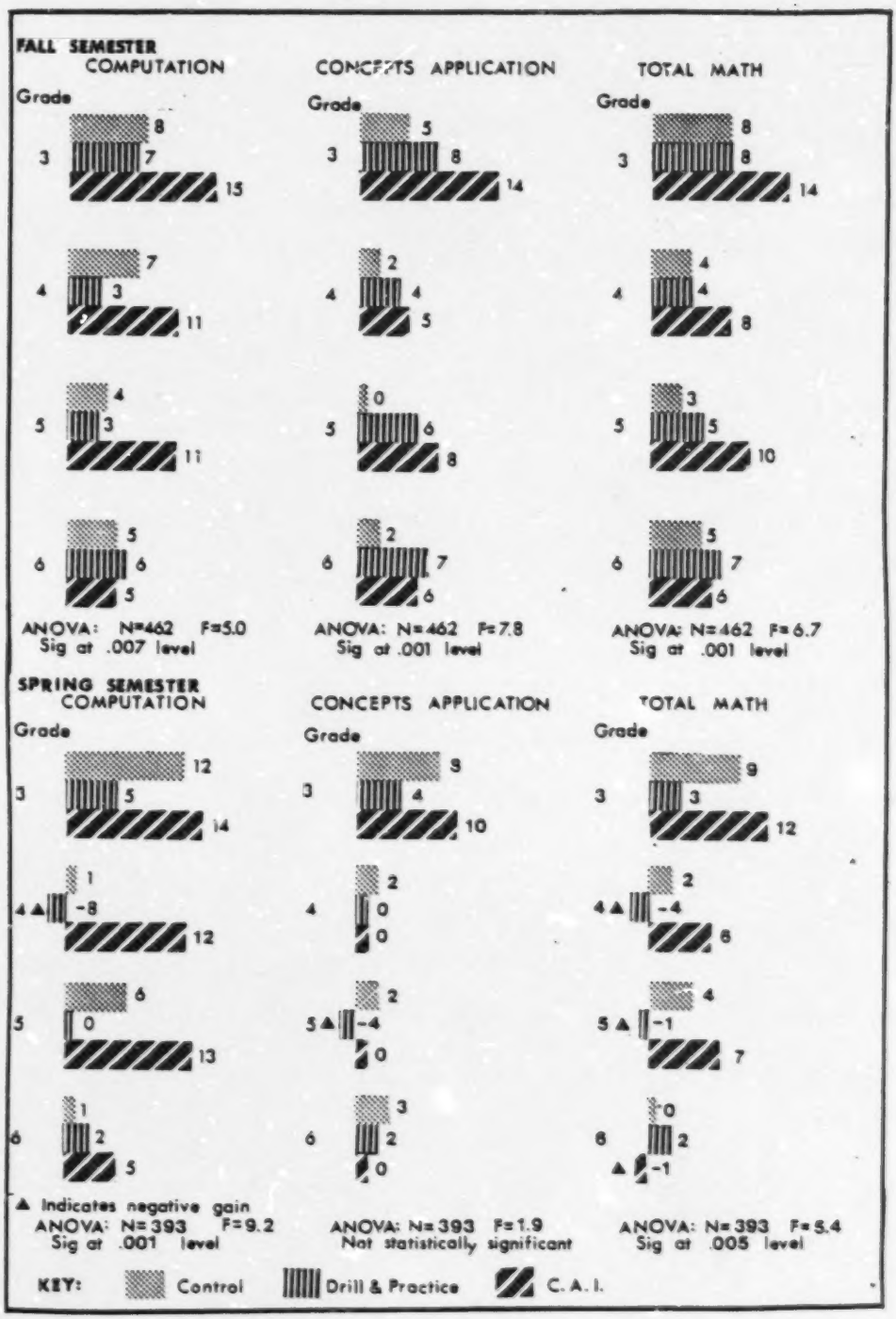

FIGURE 1. Gains in NCE scores by treatment

statistical probability that there would be this much difference between the gain scores of the three treatment groups by chance is less than one in 1000. (It is customary in educational evaluation to consider differences between groups "statistically significant" when the probability of chance difference is below the .05 level, e.g. 5 chances in 100. These levels of significance, therefore, would be considered very high.)

In the spring semester the findings represent a very similar pattern. The greatest effect of CAI appears to have been in improving computation skills where consistently this treatment produced larger gains than the other two. In third, fourth and fifth grade the total math score gain was also greater for the group in CAI than in the other treatments. And for both computation and total math the differences between scores are statistically significant at a high level.

CAI seems to have been less helpful for improving children's skills in "concepts application" than in computation and total math. This finding was also true during the fall semester, but the deficiency in this regard is more pronounced in the spring semester. As shown in Figure 1, children in the regular district math program, as represented by the control group, appeared to have made the greatest gain in concepts application, and there was no gain in this skill for three grade levels of children in the CAI program.

It would be necessary to examine the content of the computer program more carefully to explain these findings. Computation involves problems using only numbers and symbols. This type of problem would be relatively easy to present on a computer screen. Concept application problems, however, involve reading of measure-

ment devices pictorially presented, deriving information from graphs and figures, and solving "story problems" presented in two or three sentences of information. These types of problems would be more difficult to present on a computer screen and it may be that the computer program, therefore, provided little practice in this type of math problem solving.

The gain scores from the Spring semester also follow the same pattern as those in the fall in regard to the greatest gains being made by third grade children, with the program having relatively little effect at the sixth grade level. Again, it would be necessary to examine the type of skills children are expected to master at these different grade levels to see how directly the CAI program used provides practice in these skills areas. Practice in simple computation, addition, subtraction, multiplication and division probably makes up the major part of the skills the third grade child is learning in math; it may represent a relatively small part of the math curriculum at the sixth grade level.

The drill and practice program was instituted as an alternative program to CAI which provided a less expensive method of reinforcing match skills through extra practice sheets. Figure 1 shows that this program may have been counterproductive. Except at the sixth grade, which follows a different pattern than the other grades, children in the drill and practice program consistently made lower gains in computation than children in either CAI or the regular district math program. This was true both semesters.

During the Spring semester when children who had been in the CAI program first semester were put into the drill and practice materials, in a number of instances they show an actual decline in test scores.

Using the U.S. Dept. of Education "rule of thumb" standard that a gain of three NCE's represents a successful remedial effort and a gain of 7 NCE's an "exemplary" program, it will be seen that, the single semester gains in computation and in total math by children in the CAI program, except at the sixth grade level, were consistently large enough to be in the "exemplary" category. And, except at the third grade level, this was not true of either the drill and practice alternative or the regular district math program. The fact that these findings were replicated with a different group of children in the CAI program for each of the semesters reported, increases the reliability of this finding. When the pattern of gains is consistent in successive trials, it makes it less likely that the results occurred because of chance selection factors in the group of participating children.

Was CAI effective in improving math skills for children in remedial or special education groups?

In each math treatment group there were some children enrolled whose academic deficiencies made them eligible for remedial programs under either Chapter 1 or local RAP funding. There were also children who were classified as "special education" students in each treatment. Figure 2 shows the gain scores for these two groups of children.

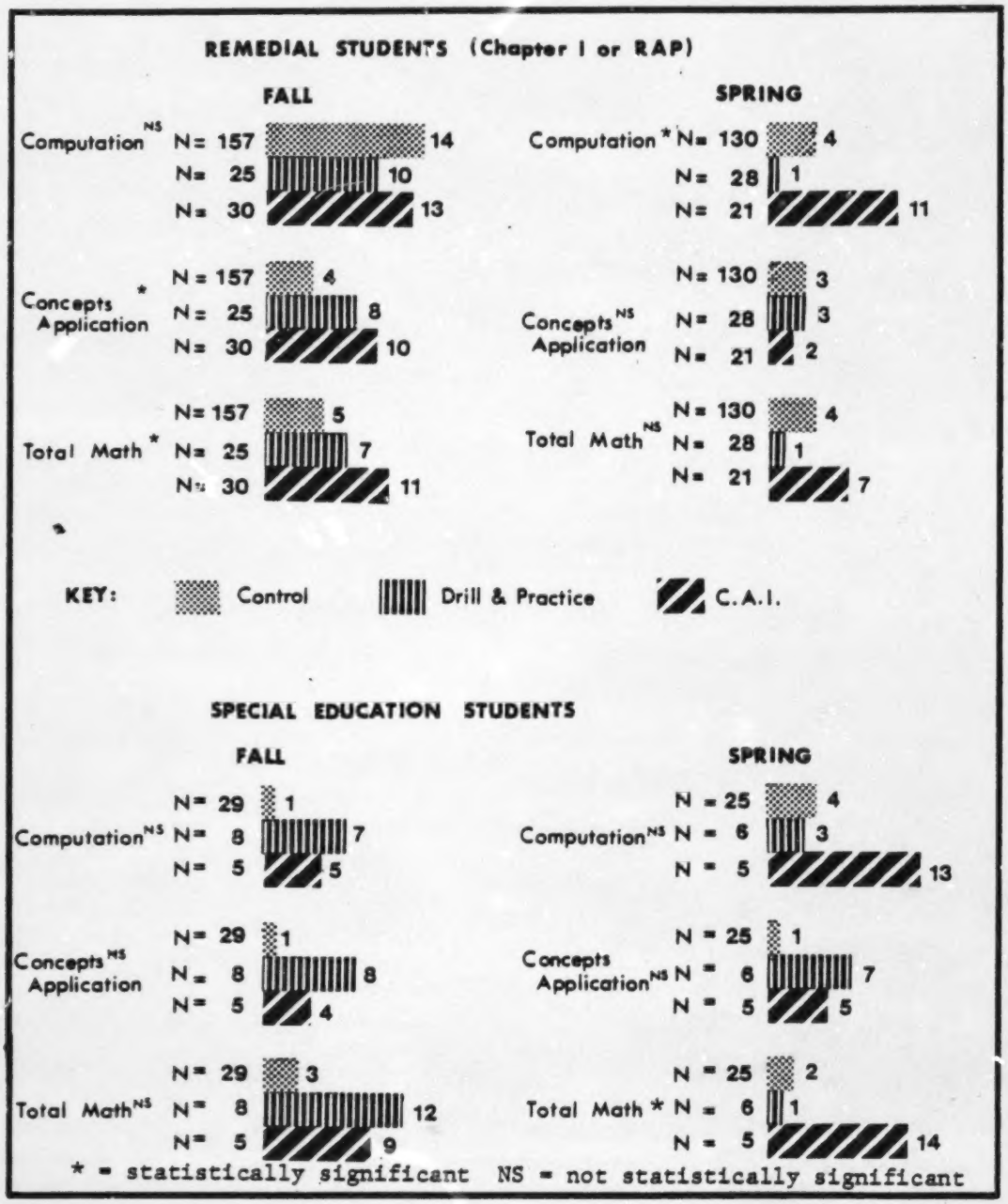

Figure 2. Gains in NCE scores in math by students eligible for Chapter 1 or RAP remedial programs, and by special education students who received computer assisted instruction or alternative math programs.

Because the numbers are much smaller, these findings are much less "stable" in terms of producing statistically reliable results presented in terms of group averages and group differences. An asterisk next to the test title indicates that the differences between groups on this subtest were "statistically significant." A designation of "NS" next to the test title indicates that the differences were not "statistically significant." These terms translate to mean that when there are greater differences between groups than there are differences in individual scores within each group, there has been found a "statistically significant" difference. If there is a great deal of variability within each group, however, even though there may appear to be large differences between the group averages, the result is not statistically significant.

Looking at Figure 2 there are four instances where the differences in scores produced a statistically significant result. This is for total math and for concepts application of remedial students during the Fall semester, and for computation scores of remedial students during the Spring semester. In the small sample of special education students, the only statistically significant difference found was for total math scores during the Spring semester. In all of these cases it should be noted that the CAI program was the one that produced the largest gains.

Was CAI effective in improving math skills for migrant children or children with limited English proficiency?

Separate analysis was done for children classified as migrant and for children shown by language tests to be dominant in another language and not proficient in English. Because many of the migrant children served by the district come from Spanish speaking families, there is some overlap between children in these two groups. The gain scores for these groups are shown in Figure 3.

Once again, because the groups are small, the effect of chance variation within the groups is often enough to render meaningless what appears to be large differences between the groups. As a result, there were found to be statistically significant differences (noted by an asterisk on the Figure) in only five out of twelve analyses. As in the previous figure, in all five cases where differences were statically significant, CAI produced the largest gains among the three treatments. Computation scores were significantly higher for children receiving CAI in both Fall and Spring semesters for both migrant and limited English proficient children. Total math was also significantly higher for migrant students during the fall semester. Even in cases where differences were not statistically significant, there is the same pattern; in nearly all cases the CAI treatment appears to have produced the largest gains. The program, therefore, appears to have been particularly effective in overcoming the language barrier present, to some degree, with these two subgroups of students.

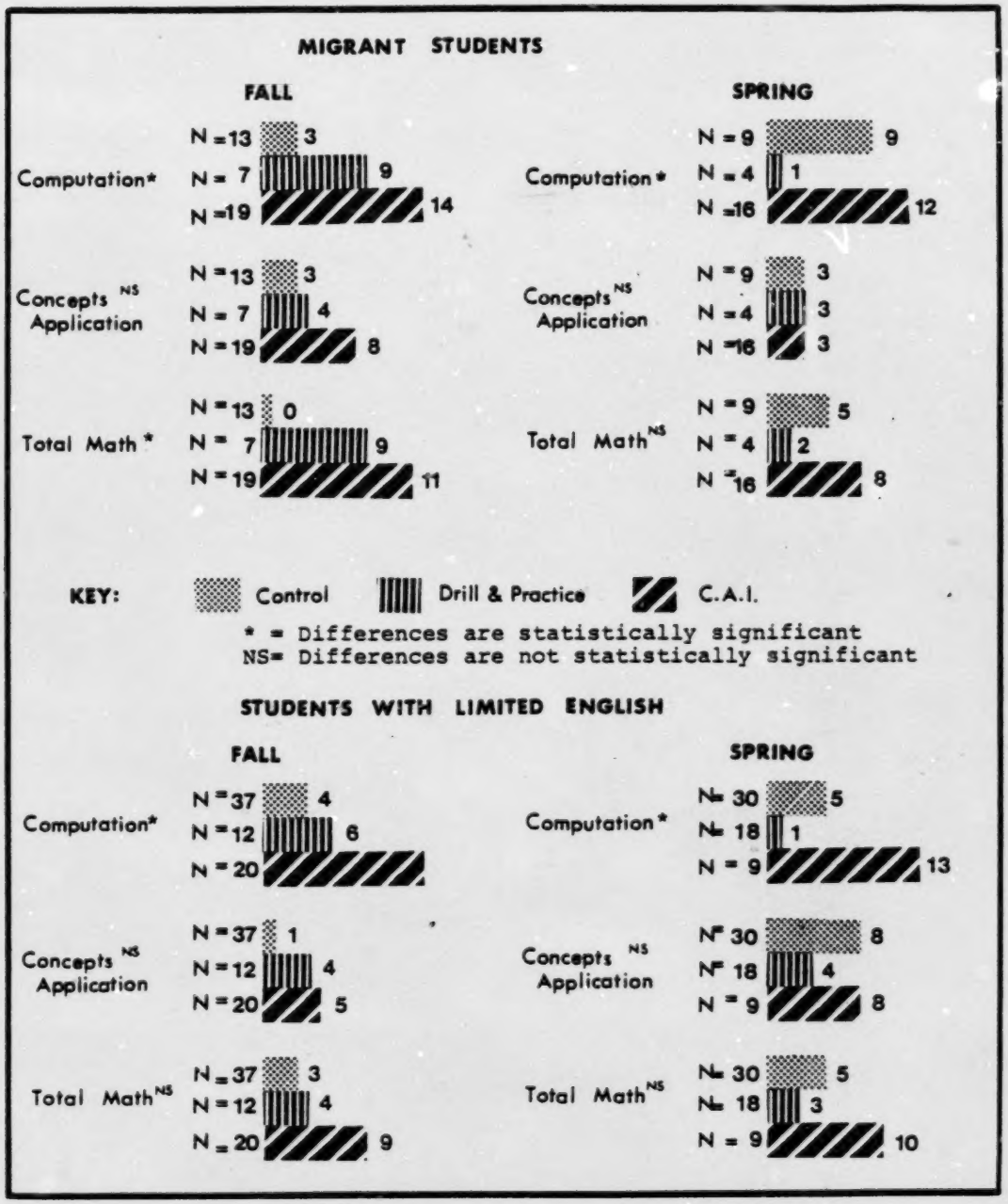

Figure 3. Gains in NCE scores in math by migrant and limited English speaking students enrolled in the computer assisted instruction or alternative math programs.

How did the different math treatments affect children's performance in relation to national norms?

Children in all three math treatment groups showed large gains in math scores during the 1982-83 school year as shown in Table 1. In the fall of the school year the average scores of all three groups were consistently below national norms (represented by an NCE of 50 for each grade level). At the end of the school year, except at the fifth grade level, average scores were as consistently above national norms for all three groups. District personnel interpreting why this might have occurred reported that there was a district wide emphasis this year on improving math achievement. There may also have been added interest in math achievement related to the CAI experiment which produced a spinoff effect through extra effort in this subject area in classrooms not involved in the experimental program. If this was the case, the experimental program has had a beneficial effect across all three schools.

Taken grade by grade, the differences between scores of the three groups were not large enough to be statistically significant. As shown in Table 1, the only groups of scores where differences were statistically significant was for computation skills, at mid-year, where the third and fourth grade children who had just completed the CAI program had markedly higher scores than the other two groups.

However, closer examination of the table confirms other reported findings. In computation, which appeared to be the type of math skill most benefitted by CAI, the highest scores in every grade level at mid-year were for group one which had just completed the CAI program. In the Spring this superiority in computation had shifted to group two, which was the group receiving CAI, and

TABLE 1

MEAN MATH SCORES BY GRADE LEVEL AND TREATMENT GROUP

Subtest Title Time of testing Treatment Group[1]	Third Grade	Fourth Grade	Fifth Grade	Sixth Grade
COMPUTATION				
Fall				
1. CAI/D&P	41.9	46.4	41.1	46.4
2. D&P/CAI	43.3	42.2	39.6	39.7
3. Control	37.7	45.1	41.4	42.8
Mid-year				
1. CAI	57.1 *	57.8 *	52.1	51.6
2. Drill & Practice	50.1	45.6	42.3	45.9
3. Control	45.6	52.1	45.7	48.3
Spring				
1. Drill & Practice	61.9	52.7	52.3	55.2
2. CAI	64.1	57.0	55.7	51.9
3. Control	59.2	54.8	50.7	48.0
CONCEPTS APPLICATION				
Fall				
1. CAI/D&P	37.1	46.8	41.4	45.7
2. D&P/CAI	40.1	48.0	37.9	47.1
3. Control	44.0	52.4	44.8	47.6
Mid-year				
1. CAI	50.6	52.1	49.1	51.2
2. Drill & Practice	47.8	51.8	44.3	52.8
3. Control	49.2	54.0	45.2	49.7
Spring				
1. Drill & Practice	54.3	52.6	46.2	53.8
2. CAI	57.0	51.4	46.0	50.4
3. Control	57.1	57.8	47.3	53.7
TOTAL MATH				
Fall				
1. CAI/D&P	40.6	46.3	40.7	47.8
2. D&P/CAI	41.8	45.0	38.1	45.3
3. Control	41.4	49.4	42.6	46.5
Mid-year				
1. CAI	54.7	54.7	50.4	53.7
2. Drill & Practice	49.5	48.6	43.2	52.5
3. Control	49.0	53.4	45.3	51.2
Spring				
1. Drill & Practice	58.0	52.2	50.2	55.7
2. CAI	60.9	54.2	50.8	50.9
3. Control	59.1	57.3	49.5	52.1

* = Differences statistically significant beyond .05 level (see note 2)

1. Note that the treatment groups 1 and 2 are listed in the fall according to the order they would receive the two treatments, and the mid-year and spring identification is by the treatment received by this group in the preceding semester.
2. Analysis of variance was done for the 3 groups of scores at each grade level. An * indicates a statistically significant difference was found between groups. For all others the differences were found to be not significant.

the highest scores at every grade level (except sixth) were once again found with the group that had just completed the CAI program.

In concepts application, an area of math skills less benefitted by CAI, the mid-year scores showed a mixed pattern, but the spring scores consistently found the control group, from the district math program, with the highest scores. It should be noted, however, that this group also started the year with the highest scores in concepts application.

For total math, which combines the skills of computation and concepts application, at mid-year the CAI group consistently had the highest scores. At the spring testing, group 2, which had just completed the CAI semester, had the highest scores in third and fifth grade. Thus, in six out of eight comparisons by class level, the children most recently in the CAI program had the highest average scores in total math achievement.

MEASURING GRADE EQUIVALENT GAINS IN CAI

In addition to the evaluation questions related to overall program effectiveness,compared to other treatments, school staff requested information on two other questions related to use of the computer program. The first was how well the grade equivalent rankings given by the computer master terminal compared to grade equivalent scores children received on standardized tests. The other was how length of time on the computer related to children's grade equivalent gains. Correlational analysis was used to answer these two questions and the findings are reported below.

How did grade equivalent ratings by the computer program compare to grade equivalent ratings from standardized tests?

The grade equivalent ratings on the standardized tests were compared to the grade equivalent ratings derived from the computer for both groups of children who participated in the CAI program during the 1982-83 school year. As shown in Table 2, below, the correlation between the scores was both high and consistent within the experimental groups, indicating that both measurement scales are making about the same assessment from a common body of skills. All of the correlations are statistically significant at the 0.000 level (e.g. probability less than one in a thousand such a result would occur simply by chance). In educational evaluation, correlations of over.60 are considered high. In this case the correlations range from .63 to .86.

Table 2

Correlation between standardized test and computer grade equivalent scores.

Group 1 CAI Fall 1982		Group 2 CAI Spring 1983	
Fall tests	Mid-year tests	Mid-year tests	Spring tests
r = 0.8610	r = 0.8491	r = 0.6275	r = 0.6559
N=73	N=73	N=53	N=53
p = 0.000	p = 0.000	p = 0.000	p = 0.000

Another way of answering this question was to compare the gains in grade equivalent scores by grade level on the standardized tests and the computer reports. This is shown in Table 3.

Table 3

Grade equivalent gains as measured by standardized tests and computer, by grade levels, for children in Pasco CAI program.

	Group 1 CAI Fall 1982		Group 2 CAI Spring 1983	
	Mean Grade Equivalent Gains		Mean Grade Equivalent Gains	
Grade	Stand. Tests	Computer	Stand. Tests	Computer
Gr. 3	10.8	9.9	10.3	6.1
Gr. 4	6.9	8.3	6.8	9.2
Gr. 5	10.2	9.4	7.2	9.2
Gr. 6	8.1	12.8	2.6	8.0
Mean of Combined Grades:	8.8	10.0	7.0	8.2
Standard Deviation	(7.7)	(3.6)	(6.3)	(2.8)

There are several things to be observed from Table 3. At the third grade level, the standardized test shows a higher gain in grade equivalent score than the gain from the computer rankings. This is consistent for both groups who participated in the CAI program. On the other hand, at the sixth grade level, the gain shown by the computer scores is much higher than the gain shown by the standardized test. The most likely meaning of this finding is that at these grade levels there is not a very good fit between the items tested on these particular standardized tests and the curriculum of skills taught by the computer program. Children at the sixth grade level may have learned a great deal of new math skills which are not being picked up by the standardized test.

The gains reported by the computer grade equivalents are far more consistent from grade level to grade level and from child

to child. This would be the result one might expect if there is a good fit between the skills taught and the measurement scale; a given amount of input should produce a more predictable gain. This is reconfirmed by looking at the standard deviations of the gains. In Group 1 the average grade equivalent gain on the standardized tests for the total group was 8.8, with a standard deviation (i.e. the variability of scores spread out around that group average) of 7.7. For the same Group the average gain in computer grade equivalent score is 10.0 with a standard deviation of only 3.6; i.e. much less variability of scores around the group average. The same is true for Group 2: for the standardized test the average gain was 7.0 with a standard deviation almost as large, 6.3; in the computer scores the gain was 8.2 with a standard deviation of only 2.8. This may explain the findings in the main body of this evaluation in which the sixth grade so often showed an opposite result from the other grades. If the test is not measuring what is being taught, no degree of effective teaching will produce a measurable result.

As shown in Table 3, overall the gains made by Group 1 were larger than the gains made at comparable grade levels by Group 2. This is also understandable. Referring back to Table 1, children at all grade levels were performing well below grade norms in the fall of 1982. Group 1, which started CAI in the fall, showed very rapid gains so that by mid-year testing children's average scores was above national norms at every grade level. The second group, however, began the CAI program at mid-year when they had much higher

starting scores (e.g. at third grade children had an NCE of 50, the national norm when they began the CAI program). The gain children made while in CAI was therefore more of a test of the program than in the first semester. It took children already performing close to grade level and pushed their scores up to a level well above national norms for that grade (e.g. third grade children ended the CAI semester with an NCE of 64, which is 14 points above the national norm). Gains are understandably harder to make when children are already performing above group norms and are approaching a ceiling level based on a normal ability curve for children of a given grade.

One other final point should be made from the findings reported in Table 3. This is that if children were showing only "expected" progress, their grade equivalent gain in half a year would be 5. Children in the CAI program exceed this standard whether measured by the computer or by standardized tests. They average far more than half a year's gain in math skills during the semester they were in the CAI program.

How were grade equivalent gains related to time on the computer?

A correlation was made between gains in grade equivalent scores and time on the computer, both for the gains measured by standardized tests and the gains by computer reports. The results are shown in Table 4 below.

Table 4

Correlation between grade equivalent gains and time on the computer based on standardized tests and computer reported grade equivalents.

Group 1 CAI Fall 1982		Group 2 CAI Spring 1983	
Time with Stand. Tests	Time with Computer G.E.	Time with Stand. Tests	Time with Computer G.E.
r = 0.2225	r = 0.2126	r = 0.1119	r = 0.2783
N=73	N=73	N=53	N=53
p = 0.029	p = 0.035	p = 0.213	p = 0.022

There is a positive correlation between time on the computer and gain in grade equivalent score on both the standardized test and the computer assigned rating for Group 1. The correlation is rather low, but it is enough to be statistically significant beyond the .05 level in both cases. For the second Group there is again a positive correlation between time on the computer and gains in grade equivalent scores. In this case the correlation is statistically significant only for the computer ratings of grade equivalent gains.

Summary

The gains made by children in CAI were consistently higher than that established by the U.S. Department of Education as "educationally significant." The gains were also greater than that of comparable children enrolled in either the district's regular math curriculum, or children receiving comparable time in paper and pencil drill and practice exercises in math. The differences in gains between CAI and these other treatment groups were

large enough to be statistically significant for computation, concepts application and total math for the group receiving CAI during the first semester, and for computation and total math for the group receiving CAI during the second semester.

The CAI program seemed to produce the greatest gains in computation skills and to be less effective in developing concepts application skills, as measured by these standardized tests.

The younger children, third grade, seemed to benefit most from the program and the older children in sixth grade to benefit the least. However, this difference may be exaggerated by a poor fit between the tests used to measure gains at these grade levels and the array of math skills taught through the CAI program.

Children enrolled in special education performed better in CAI than in the regular math curriculum. Children in district remedial programs (Chapter 1 or RAP) also tended to make greater gains through CAI than through the regular district program. For these two groups all statistically significant differences favored the CAI treatment.

The CAI treatment seems to have been particularly helpful in overcoming the language barrier for children with limited English. Special analysis for children of limited English proficiency, and for migrant children, many of whom are from Spanish speaking homes, showed consistently greater gains in computation and total math for children receiving CAI, and all statistically significant differences in math gains favored children in CAI.

During the 1982-83 school year, all three math treatments brought children from below to above grade norms in total math. At the end of the year the control group in the regular district math program had made the greatest gain in concepts application, the CAI groups the greatest gains in computation and total math.

There was a positive correlation between length of time on the computer and gains; e.g. the more time on computer the greater the gains.

In summary, the CAI program, as implemented by Pasco Public Schools, is a very powerful program to accelerate children's learning of basic math skills. The replication of findings with two groups of children reported here strengthens this general conclusion.